Naval Research Laboratory

Washington, DC 20375-5320

NRL/FR/5750--04-10,080

ShipIR Model Validation Using NATO SIMVEX Experiment Results

DOUG FRAEDRICH
CRAIG MILLER

Tactical Electronic Warfare Division
Advanced Techniques Branch

ESPEN STARK
LARS TRYGVE HEEN

Norwegian Defence Research Establishment
Kjeller, Norway

May 31, 2004

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
31-05-2004

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
October 2001-December 2003

4. TITLE AND SUBTITLE
ShipIR Model Validation Using NATO SIMVEX Experiment Results

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
PE 0603790N

6. AUTHOR(S)
Doug Fraedrich, Craig Miller, Espen Stark*, and Lars Trygve Heen*

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
Code 5750
4555 Overlook Avenue, SW
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL/FR/5750--04-10,080

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Chief of Naval Resesarch (ONR 313)
1725 Jefferson Davis Highway
Arlington, VA 22217-5660

10. SPONSOR / MONITOR'S ACRONYM(S)
ONR

11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
*Norwegian Defense Establishment
Kjeller, Norway

14. ABSTRACT

An infrared (IR) field trial has been conducted by a NATO science panel on IR ship signatures, TG-16. This trial was planned, designed, and executed to validate predictive IR ship signature simulations. Details of the trial were dictated by a validation methodology that exploits the concept of "experimental precision." Two government defense laboratories, the Norwegian Defence Research Establishment and the U.S. Naval Research Laboratory, have used these trial data to perform a validation analysis on the ShipIR IR signature code. This analysis quantifies prediction accuracy of the current versions of the code and identifies specific portions of the code that need to be upgraded to improve prediction accuracy.

15. SUBJECT TERMS
Ship signature, Infrared, Model validatioin

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Douglas S. Fraedrich
a. REPORT	b. ABSTRACT	c. THIS PAGE	UL	13	19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified			(202) 404-3810

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

CONTENTS

SHIPIR MODEL VALIDATION USING NATO SIMVEX EXPERIMENT RESULTS

INTRODUCTION

ShipIR is an infrared (IR) ship signature prediction model that has been adopted by both the North Atlantic Treaty Organization (NATO) and the United States Navy as a common tool for predicting the IR signature of navy ships in a maritime background. Since 1993, when the model was accepted as the NATO standard, a series of model upgrades have been implemented that have been guided by comparison with field data. The results of a preliminary validation of a previous version of the model (Version 2.5) were described in an earlier paper [1]. A NATO science panel, Task Group-16 (TG-16), which has the charter to research IR ship signature issues, has recently executed a field trial to validate the latest versions of the ShipIR model (Versions 2.9 and 3.0).

ShipIR is a model that simulates the IR radiance of both ship targets and the maritime background. It is composed of several major submodels: sky radiance, sea radiance, plume radiance, ship temperature, and ship surface emission/reflection.

The current study describes the results of a validation exercise based on data from the NATO SIMVEX trial (Ship Infrared Model Validation EXperiment). This trial was executed specifically for IR model validation; its design was based on the collective experience of the NATO TG-16 countries from each of their individual validation efforts. The participating countries (along with the specific organizations) were: Canada – Defence Research and Development Canada (DRDC) and Davis Engineering; Norway - Norwegian Defence Research Establishment (FFI); United States - Naval Research Laboratory (NRL); The Netherlands - TNO Physics and Electronics Laboratory (FEL); Italy - Mariteleradar; France - Centre Technique des Systemes Navals (CTSN); Denmark - Danish Defence Research Establishment (DDRE); and Poland – Naval University of Gdynia.

This paper summarizes the results of two national laboratories: the Naval Research Laboratory (United States) and the Norwegian Defence Research Establishment (Norway). These two countries were the first to complete their analyses. Comparisons between measured and predicted signature are shown for both versions, V2.9 and V3.0, of ShipIR. The prediction accuracy of the model is quantified in both the mid-wave IR (MWIR) (3-5 μm) and long-wave IR (LWIR) (8-12 μm) spectral bands, and errors have been traced to individual submodels for subsequent improvement.

DESCRIPTION OF THE SHIPIR MODEL

ShipIR [1] is a model that simulates the steady-state IR radiance of ship targets in the maritime background (Fig. 1). This model was developed by Davis Engineering and was originally funded by the Canadian Department of National Defence, through the Defence Research Establishment Valcartier. The model is composed of several major submodels: an IR sky radiance and propagation model; a sea reflectance model; a three-dimensional (3D) ship surface model, which enables the modeling of complex ship geometries; a

Manuscript approved January 20, 2004.

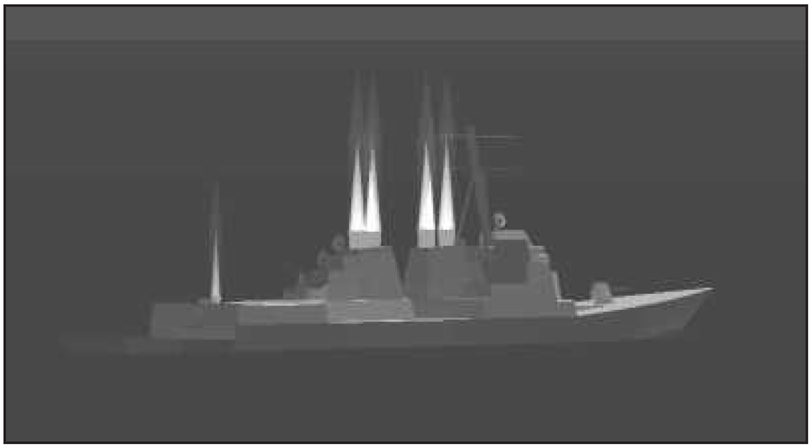

Fig. 1—Typical output of model showing the various elements of sky, sea, ship, and plume

heat transfer model; a surface radiance model, capable of simulating multibounce reflections; and an IR plume model, which supports predictions of both diesel and gas turbine plume radiance profiles (Fig. 2).

The 3D ship surface model imports a DXF-formatted CAD file. MODTRAN 4 is used to compute the radiance distribution of the sky. Sea radiance is computed using an improved version of a sea reflectance model [2]. Heat transfer is modeled by accounting for conductive, radiative, and convective terms. To estimate the local convection over the ship, each ship facet is classified as either in tangential or separated flow, with separate correlation relations (Nusselt number vs Reynolds number) being used in each flow regime. View-dependent facet radiance is determined by combining emitted and reflected components. The surface reflectance is modeled by a multiparameter Gaussian bidirectional reflectance distribution function (BRDF) superimposed on a diffuse reflectance component. Many-order diffuse multibounce terms are computed using a standard "radiosity approximation." A single specular bounce term is computed and added to the diffuse result. Plume radiance is computed using an empirical ship plume flowfield combined with NASA IR band emission models for CO_2, H_2O, CO, and soot.

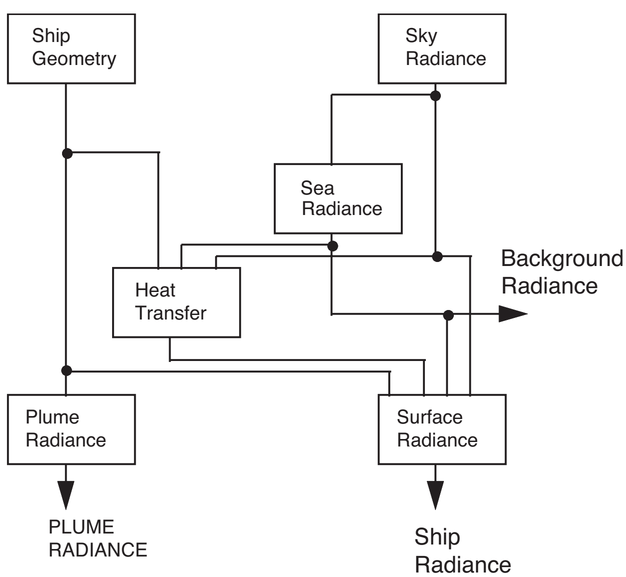

Fig. 2 — Block diagram of ShipIR illustrating submodels

DESCRIPTION OF THE NATO SIMVEX TRIAL

The measurements were performed at the Cape Scott test facility at Osborne Head (near Halifax), Canada. The test was conducted from 6-21 September 2001. This includes the time to mount and disassemble the trial equipment. As previously stated, the participating countries were Canada, Denmark, France, Italy, The Netherlands, Norway, Poland, and the United States. Figure 3 shows the facility with all the different measurement teams in position.

Each country (except Canada) contributed different IR cameras for the measurement. Canada contributed the test site and the ship to be measured; the ship used was the CFAV *Quest* (Fig. 4), which is operated by DRDC for various acoustic and oceanographic studies. Canada equipped the ship with a meteorology station and approximately 140 sensors for ship surface temperature measurements. Several additional thermocouples were mounted on the ship's exhaust uptakes, and a rented gas analyzer was used to measure the plume constituents. Norway provided a Global Positioning System (GPS) capable of transmitting the ship position to the shore site and a shore-based weather station. At the shore site, the ship position was distributed to all measurement stations through the local area network at Osborne Head.

A total of 40 runs were executed during the trial: four radiometric accuracy runs, five plume-only runs, one DRDC panel run, and 30 full-ship runs. For the five plume runs, the *Quest* was configured to sail at different power settings for the diesel engines: single engine at 50% power, single engine at 100% power, and two engines at 100% power. The two-engine runs were designed to test the accuracy of the blending in the plume rendering algorithm.

Four different full-ship runs were performed during the trial. The four full-ship run types were: Type A - heading 270°, measure shaded starboard side; Type B - heading 360°, measure sunlit port side; Type C - heading 032°, measure sunlit port side; and Type D - heading 032°, measure port side at night. Of the 30 full-ship runs, there were four Type A, five Type B, eight Type C, and 13 Type D. The specific headings were determined by the local shore geometry and solar elevation/azimuth angles (Fig. 5). D runs were performed at night along the same heading as the C runs.

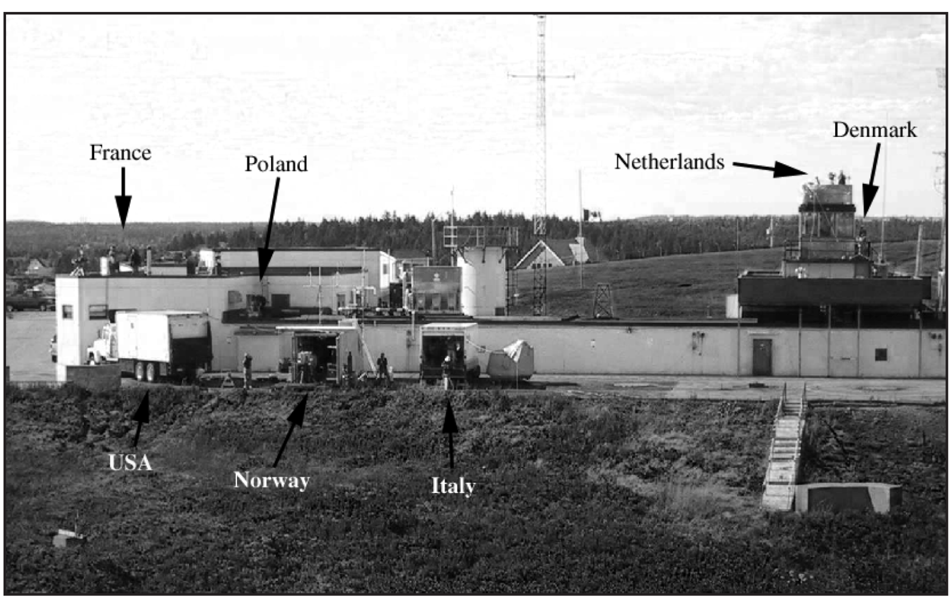

Fig. 3 — Test site at Osborne Head, Canada

Fig. 4 — Test ship CFAV *Quest*

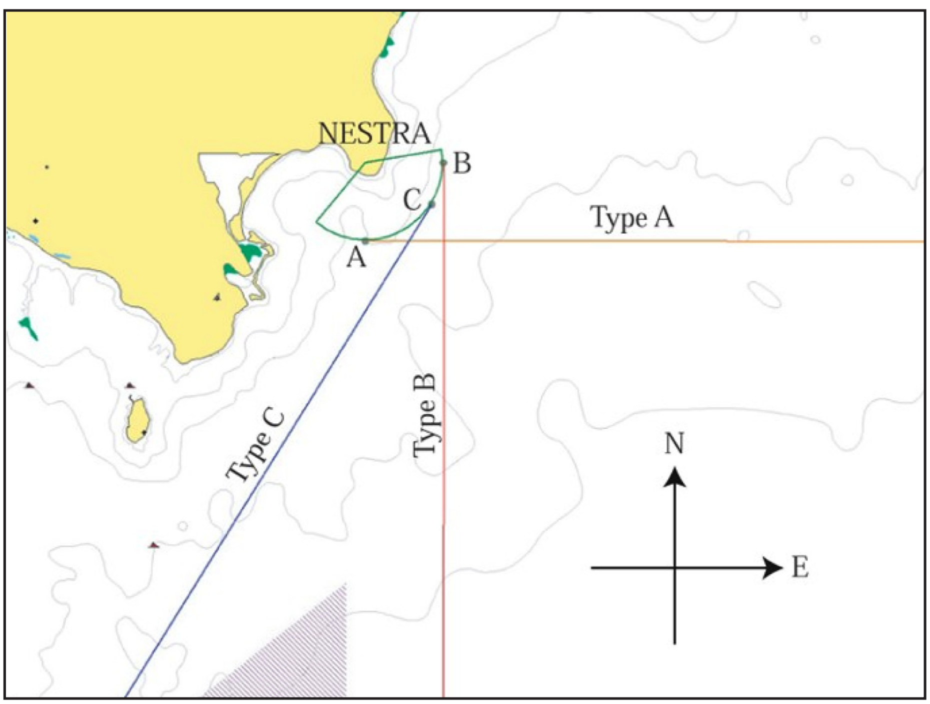

Fig. 5 — Test site (NESTRA) and run Types A, B, and C

For all full-ship runs, the ship sailed with a constant speed and heading for 30 minutes prior to measurement. The measurements were made at the point of closest approach. This point is on a 1-km arc from the measurement station and is marked on the map in Fig. 5. A constant speed of 10 kts was used, which required the ship to start 5 nmi from the point of closest approach. This allowed the ship to reach thermal equilibrium before each measurement, which is an important criterion for the validation of ShipIR.

After the second day of ship runs, the U.S. team installed their radiometers in a commercial helicopter and began performing both plume and full-ship runs from the airborne platform. This enabled the measurement of more aspect angles for the full-ship runs.

INSTRUMENTATION

Several types of instruments were involved in the trial: meteorological, temperature measurement, exhaust gas measurement, and infrared radiometric (imaging radiometers and a spectroradiometer). Since this trial was conducted to validate an infrared signature model, the primary instruments were the IR imaging radiometers. Table 1 lists three of these instruments. Spectral response, extensive calibrations, and error analyses were performed on each system. Measurement accuracy for each instrument was determined through pre-trial joint calibration/accuracy tests and post-trial error analyses [3, 4].

Table 1 — Primary Infrared Instrumentation

Country	Instrument	Image Size	Band (μm)	Accuracy (%)
Norway	Inframetrics MilCam	256 × 256	3.4-5.05	11
U.S.	Indigo Merlin	320 × 240	3.4-5.0	9
U.S.	Agema 880 LWB	140 × 140	8.0-11.5	8

VALIDATION ANALYSIS METHODOLOGY

Many of the countries have previously performed validation studies of ShipIR using their own domestic ships. Lessons learned from these many previous trials, as well as methodological procedures developed by NRL [5], were used in designing this trial.

The overall objective of this validation effort is to quantify the IR signature prediction accuracy of the ShipIR model. Measured signature values of the *Quest* were compared to predicted signature values, using the appropriate ship operational and environmental parameters for each run. Several methodological principles were used that distinguished this study from previous efforts.

- High-fidelity ship representation: The geometrical representation of the *Quest* was composed of approximately 6400 facets and 1400 isothermal plates, which is greater detail than had been used in previous validation studies. A total of 13 different radiative materials were modeled (different paints, windows, canvas, etc.), and detailed reflectance measurements were made of witness samples of all major materials.

- Error diagnosis: Additional data were collected on parameters that were neither inputs nor outputs to the simulation. Ship surface temperature, direct solar irradiance, and absolute sky/sea radiance measurements were made. These additional data enable diagnosis of the source of prediction error at the submodel level. Such information is critical in determining which part of the model to improve to yield better prediction accuracy for future releases.

- IR radiometric accuracy trials: Before the trial, all measurement teams performed measurements on a 4 × 4-ft test panel of known temperature and emissivity. This enabled all measurement teams to

estimate their measurement accuracy and improve calibration, data collection, and data reduction procedures as necessary.

- Accurate and comprehensive measurements of input data: Meteorological measurements were made from the shore, the ship, and from a buoy. In addition, nearby weather balloon measurements were used to model the proper vertical profile of the atmosphere.

The concept of "experimental precision" [6] combines the last two items: measurement accuracy of model outputs and inputs. The precision of the experiment, for purposes of validation, is the root-sum-of-squares (RSS) sum of the measurement accuracy of the model output and the uncertainty of model output caused by errors in all of the inputs (derived via sensitivity analysis). This precision represents a lower limit as to how accurate a model can be shown to be by using the results of the validation experiment. For this trial, the approximate level of experimental precision is 15-20% for daytime runs and 20-25% for nighttime runs.

Since ShipIR models only 0% and 100% cloud cover, all partial cloud cover runs (non-plume) were excluded from the validation analysis. Thus, for validation of the full-ship model, 17 runs were available for analysis: four Type A, two Type B, four Type C, and seven Type D.

Temperature and IR radiant intensity comparisons were made by both Norway and the United States.

VALIDATION RESULTS

Norway's analysis focuses on V2.9, MWIR band, and clear-sky runs for both day and night [7]. Temperature analyses were performed for nine runs, and IR analyses were performed for 10 runs. Norway performed spectral measurements during the plume runs (and also made spectral measurements of the ship and background); these results will be presented in a separate paper [8]. The U.S. team's analysis focused on V3.0, both MWIR and LWIR bands, and daytime runs (both clear and cloudy). The U.S. team was not able to make nighttime measurements because the helicopter could fly only during the day. The only substantive difference between V2.9 and V3.0 is improvements made to the plume emission model. Since the plume accounted for approximately 1% of the full-ship intensity, comparisons of full-ship results between the two versions are valid.

Norway Results

Figure 6 compares ship surface temperature. The runs are grouped into day runs (8, 14, 19, and 32) and night runs (15, 16, 20, 21, and 34). For the combined set of day and night runs, the overall root-mean-square (RMS) temperature prediction error is 2.6 °C. These results indicate a clear pattern that the model overpredicts during the day (average of +1.8 °C) and underpredicts at night (average of 3.0 °C).

Figure 7 shows an analogous comparison of IR contrast intensity; the day runs were overpredicted (by 36%) and the night runs were underpredicted (by 63%). While the day runs indicate a reasonable predictive capability of the model, the night runs clearly illustrate significant modeling errors. One might expect large percentage errors for the case where a small contrast is the result of the difference of two radiance values very close to each other. Analysis of the nighttime measured data reveals that the average contrast radiance difference was 8 $\mu W/cm^2/sr$, which is equivalent to an effective temperature difference of about 1 °C. Therefore, to achieve a 50% prediction error for "effective contrast temperature," a prediction error of 0.5 °C is needed for both target and background. This is quite challenging to achieve.

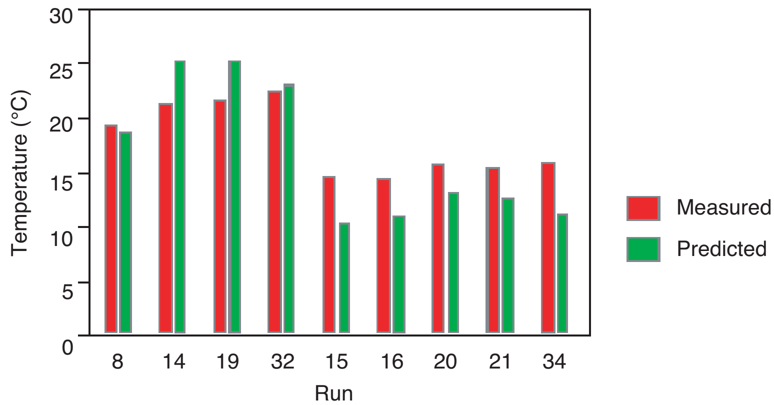

Fig. 6 — Predicted vs measured hull temperatures from Norway analysis

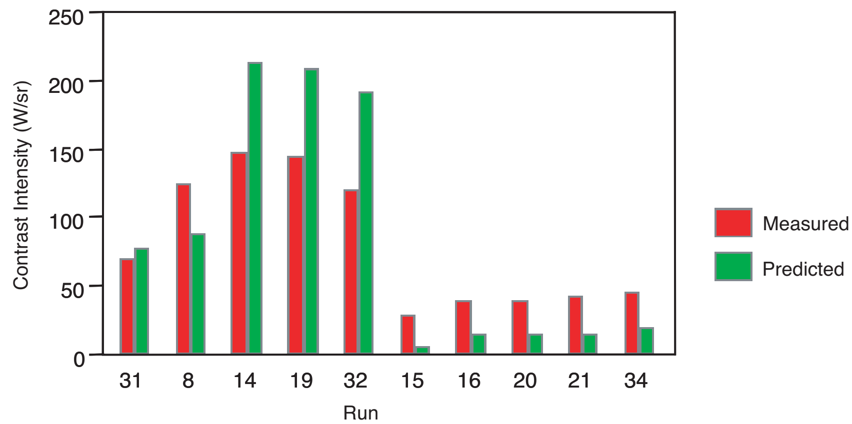

Fig. 7 — Predicted vs measured MWIR intensity from Norway analysis

For the case of the SIMVEX night runs, a sensitivity analysis was performed to obtain an estimate of "experimental precision." This indicated that the predicted signature was very sensitive to whether shore-based or ship-based meteorological data were used as input. Thus, the experimental precision of the night runs was significantly worse than that of the day runs

United States Results

As previously stated, the U.S. analysis focused on daytime runs. Most of these measurements were made from a helicopter, which enabled collection of data from all four cardinal aspects of the ship (bow, starboard, stern, and port). Also, these data were taken at a higher elevation angle, typically 2 deg, as compared to 1.4 deg from the shore site.

Figure 8 shows prediction error of surface temperature, using thermocouple data from bow, starboard, stern, and port facets. The RMS temperature prediction error for this dataset is 3.0 °C. For each run, the relative wind was computed and each side of the ship was categorized as to whether it was in stagnation, tangential, or separated convection flow. The difference between predicted and measured temperature is plotted, grouped by flow type. Mean and standard deviations are overlaid for illustrative purposes. These data show that the current convection model works relatively well for tangential flow, but overestimates temperature for both stagnation and separated flows. Currently, the convection algorithm models the stagnation case as though it were tangential. For all these data points, the air temperature was less than the predicted ship surface temperature. This indicates that convection is being underestimated for these two flow regimes. The possibility of radiative and conductive errors was eliminated by appropriate analysis.

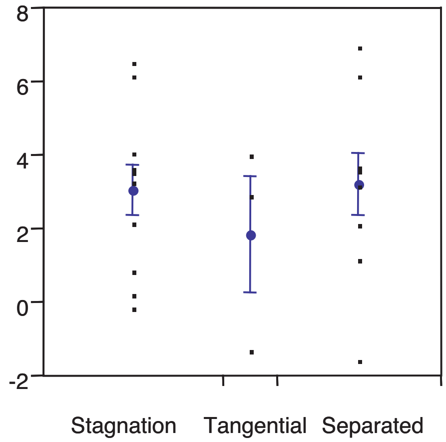

Fig. 8 — Prediction error of surface temperature as partitioned by flow regime

Figure 9 shows an overlay plot of measured and predicted ship signatures in both MWIR and LWIR bands. All of these points are day runs: clear skies for cases 1-14, and overcast skies for cases 15-18. With the exception of the first two points (which were broadside aspects measured from shore), the points are in groups of four: four runs with four measured aspects per run. The overall prediction accuracy is 34% in the MWIR band and 18% in the LWIR band.

The LWIR results are within the precision of the experiment for both clear and overcast skies. Therefore, the model is essentially as good as the measurements (with the exception of cases 1, 2, and 10, which we believe are due to temperature underestimation). The MWIR results also suffer from temperature underestimation problems (cases 1, 2, and 10) but also show significant model underestimation for cases 4, 8, and 12. All of these cases correspond to the shaded broadside aspect of the ship, measured under clear skies. Subsequent analysis of the thermocouple time histories indicate that for these runs, the shaded side of the ship was not in thermal equilibrium when measured [9]. This indicates a flaw in the design of the validation trial for shaded aspects of the ship, and is not indicative of poor model predictive capability. Figure 10 shows representative measured and predicted MWIR images.

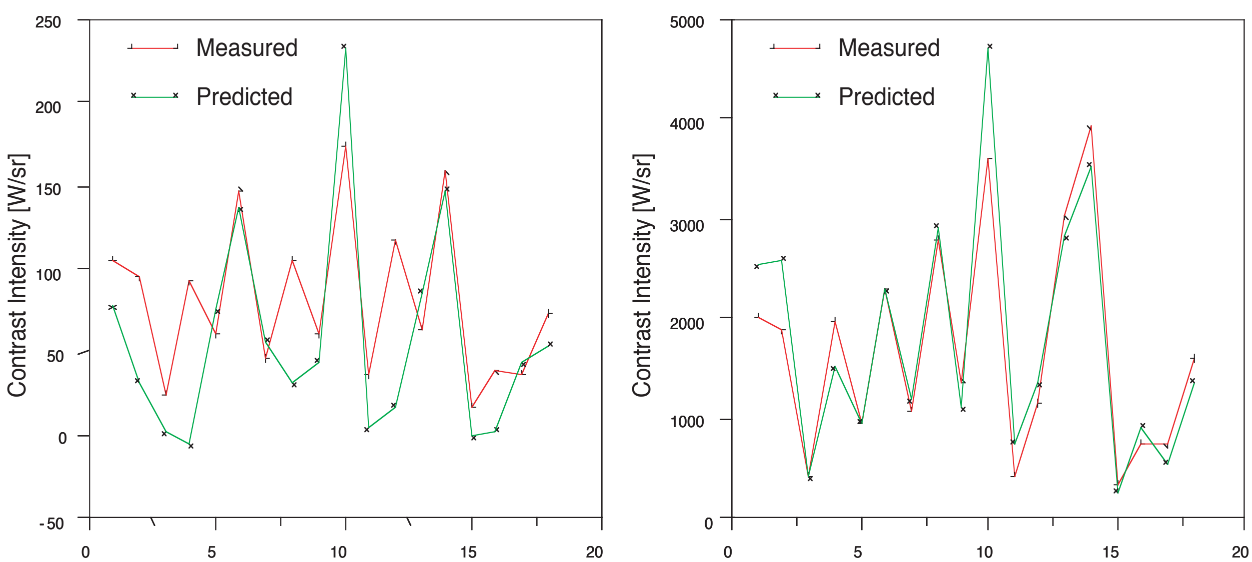

Fig. 9 — Predicted vs measured ship intensity from U.S. analysis

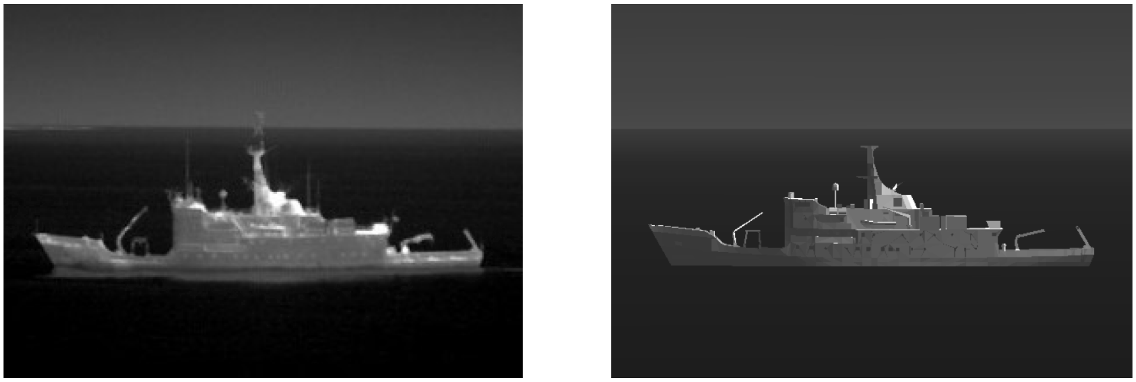

Fig. 10 — MWIR ship images

IMPLICATIONS FOR FUTURE MODEL UPGRADES

Based on these two datasets/analyses, several conclusions can be drawn. LWIR daytime prediction accuracy is within the experimental precision, given measurement accuracy and sensitivity to uncertainties in inputs. Where the two analyses overlap (MWIR daytime), the results are consistent and indicate a prediction accuracy in the 35% range. While not as good as the LWIR band, this is still adequate for most applications. Nighttime MWIR prediction accuracy is on the order of 60-65%. Although the experimental precision for the night runs is quite poor, there are undoubtedly significant prediction errors.

Subsequent analysis indicates that the majority of the observed prediction error can be accounted for by errors in ship surface temperature prediction. The Norwegian FFI dataset indicates that the heat transfer model may be underestimating convection for both day and night. The NRL temperature dataset shows that, during the day, the temperature error depends on the flow conditions at the facet in question. NRL has performed subsequent research into better Nusselt number correlation equations for stagnation and separated flows for incorporation into Version 3.1 of the model. These new correlations will result in generally higher levels of convection, which will tend to lower daytime temperatures and raise nighttime temperatures (since the air temperature at night is often above the predicted ship facet temperature).

Also, an analysis of the FFI night runs indicate that for four of these runs (20, 21, 26, and 27) the ShipIR heat transfer model predicts surface temperatures that are below the dewpoint, which is not physically reasonable. This is because condensation is not included in the current model. Version 3.1 will incorporate film condensation effects [10] into the heat transfer model. This modification will tend to increase nighttime temperatures when condensation is a factor (high dewpoint and net radiative cooling). Both of these upgrades to the heat transfer model should result in better prediction accuracy for this dataset.

CONCLUSIONS/SUMMARY

An IR field trial has been conducted by a NATO science panel on IR ship signatures, TG-16. This trial was planned, designed, and executed specifically to validate predictive IR ship signature models. The details of the trial were dictated by a thoughtful validation methodology that exploits the concept of experimental precision.

Two governmental defense laboratories, the Norwegian Defence Research Establishment and the U.S. Naval Research Laboratory, have used these trial data to perform a validation analysis on the ShipIR IR signature code. This analysis quantifies prediction accuracy for the current versions of the code and identifies specific portions of the code that need to be upgraded to improve prediction accuracy.

The prediction accuracy in the LWIR is quite good and is of similar magnitude as the experimental precision. In the MWIR, the model tends to overpredict during the day and underpredict at night. The cause of this has been traced to the heat transfer model, and specific upgrades have been identified that will result in improved prediction accuracy.

ACKNOWLEDGMENTS

The authors acknowledge the contributions of Francoise Reid at DRDC, Canada, for hosting the trial and the other NATO TG-16 participants. Also, the authors are grateful to Dr. David Vaitekunas, of Davis Engineering, for reviewing this manuscript and countless insightful discussions on this topic.

REFERENCES

1. D.A. Vaitekunas and D.S. Fraedrich, "Validation of the NATO-Standard Ship Signature Model (SHIPIR)," *SPIE* **3699**, 103-113 (1999).

2. M.D. Mermelstein, E.P. Shettle, E.H. Takken, and R.G. Priest, "Infrared Radiance and Solar Glint at the Ocean-Sky Horizon," *Appl. Opt*. **33**(25), 6022-6034 (1994).

3. International Organization for Standardization, *Guide to the Expression of Uncertainty in Measurements,* ISBN 92-67-10188-9, ISO, Geneva, 1993.

4. D. Fraedrich, "Methods in Calibration and Error Analysis for Infrared Imaging Radiometers," *Opt. Eng*. **30**, 1764-1770 (1991).

5. D. Fraedrich and A. Goldberg, "A Methodological Framework for the Validation of Predictive Simulations," *Eur. J. Oper. Res*. **124**, 55-61 (2000).

6. D. Fraedrich and R. Gover, "Statistical Uncertainty Analysis as a Tool for Model Validation," *Proceedings of the Third International Symposium on Sensitivity Analysis of Model Output,* Madrid, Spain, 2001.

7. E. Stark, L.T. Heen, and K. Wikan, "*SIMVEX 2001 Trial - Radiant Intensity Contrast,*" FFI Report 2002/02568, Kjeller, Norway, June 2002.

8. E. Brendhagen and L.T. Heen, "ShipIR Model Validation Using Spectral Measurement Results from the NATO SIMVEX Trial," to be published in *Targets and Backgrounds IX: Characterization and Representation, Proceedings*, *SPIE* **5075**, Orlando, Florida, 2003

9. D. Vaitekunas, "Preliminary Assessment of the SIMVEX Database," presented at the *Fall 2002 NATO ShipIR Workshop*, Quebec, October 2002.

10. W. Kays and M. Crawford, *Convective Heat and Mass Transfer*, Third Edition (McGraw-Hill, New York, 1993).